

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/SE05/000427

International filing date: 23 March 2005 (23.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: SE
Number: 0400821-5
Filing date: 30 March 2004 (30.03.2004)

Date of receipt at the International Bureau: 26 April 2005 (26.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

**PATENT- OCH
REGISTRERINGSVERKET
SWEDEN**

Postadress/Adress
Box 5055
S-102 42 STOCKHOLM

Telefon/Phone
+46 8 782 25 00
Vx 06-782 25 00

Telex
17978
PATOREG S

Teletax
+46 8 666 02 86
08-666 02 86

ARRANGEMENT AND METHOD FOR DETERMINING POSITIONS OF THE TEATS OF A MILKING ANIMAL

TECHNICAL FIELD OF THE INVENTION

The present invention generally relates to dairy farm robot
5 milking and to automatic attachment of teat cups related
thereto.

DESCRIPTION OF RELATED ART AND BACKGROUND OF THE INVENTION

In a known milking system, wherein teat cups are automatically
attached to the teats of a milking animal to be milked, a robot
10 arm with a gripper is provided to grip and hold teat cups during
the attachment of the teat cups. A laser emitting laser light
and a video camera provided to register laser light as reflected
from the teats of the milking animal are mounted on the robot
arm. By aid of a method known as laser triangulation, the
15 positions of the teats can be calculated. The movement of the
robot arm can then be controlled in response to the calculated
positions to be capable of finding the teats for teat cup
attachments.

A drawback of such a milking system is that the camera, while
20 being moved close to the milking animal, is exposed to dirt and
possibly physical contact with the milking animal since the
milking animal can make sudden movements. Further, the video
camera can only be in active mode to seek for the teats when the
robot arm already has collected a teat cup and initiated a
25 movement towards the teats since the camera is fixedly mounted
on the robot arm. Still further, the video camera occupies a
considerable space on the robot arm, which may limit the use of
the system.

and teats may not be visible to the cameras due to self-occlusion.

Another problem arises since each cow's physiology differs; the udders of the cows may be located at quite different locations, which put limitations on the positions of the cameras. For instance, the area where the teats most likely are has been found by empirical studies to be at least 500 x 600 x 480 mm³ large.

Another problem arises since both color and texture of the teats are similar to the surface of the udder, which means that teat detection will be an arduous task: the contrast is low and color filters are of no use. The situation is even more complicated by the fact that the size, shape, color structure, morphological structure and texture may vary quite much from animal to animal.

Accordingly, it is an object of the present invention to provide an arrangement and a method for determining positions of the teats of a milking animal in a milking system comprising a robot arm for automatically attaching teat cups to the teats of a milking animal when being located in a position to be milked, and a control device for controlling the movement of said robot arm based on dynamically determined positions of the teats of the milking animal, which arrangement and method are based on stereo vision and solve at least some of the problems of the prior art as set forward above.

It is in this respect a particular object of the invention to provide such an arrangement and such a method, which use a stereoscopic calculation method based on repeatedly recorded pairs of images, wherein the arrangement can be used for multiple purposes.

It is a further object of the invention to provide such an arrangement and such a method, which are robust, effective, fast, precise, accurate, reliable, safe, easy to use, and of low cost.

- 5 It is still a further object of the invention to provide such an arrangement and such a method, which are capable of obtaining a very high number of correct teat cup attachments.

These objects among others are, according to the present invention, attained by arrangements and methods as claimed in
10 the appended patent claims.

Further characteristics of the invention and advantages thereof, will be evident from the following detailed description of preferred embodiments of the present invention given hereinafter and the accompanying Figs. 1-8, which are given by way of
15 illustration only and are thus not limitative of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1-4 display each schematically an automated milking station including an arrangement for determining positions of
20 the teats of a milking animal according to a respective preferred embodiment of the present invention. Figs. 1, 3 and 4 are perspective views, whether Fig. 2 is a top view. The milking station is only schematically outlined in Figs. 2-4.

Fig. 5 is a schematic diagram illustrating various image
25 processing methods, among them methods included in further preferred embodiments of the present invention.

Figs. 6-7 are each a pair of images as taken by a camera pair as comprised in any of the arrangements shown in Figs. 1-4 and illustrate yet a preferred embodiment of the invention.

5 Fig. 8 is a side view of the lower part of a teat of a milking animal illustrating self-occlusion, a phenomenon which is compensated for by still a preferred embodiment of the invention.

DETAILED DESCRIPTION OF EMBODIMENTS

10 The outline of this description is as follows. Firstly, a milking system wherein arrangements and methods for determining positions of the teats of a cow according to the present invention may be implemented is overviewed. Thereafter, various camera arrangements as used in the invention are considered. The following four sections are mainly devoted to image processing.

15 The first of these sections deals with image processing in general and teat detection in particular. Thereafter, stereoscopic calculation methods for determining teat positions are considered. The following section deals with various calibration methods used in the invention. Finally, image

20 processing methods for obtaining further functionality are disclosed.

1. The milking system

25 In Fig. 1 is shown a milking system or station 3 arranged for voluntary milking of freely walking animals such as e.g. cows, i.e. the animals enter the milking station 3 in order to be milked on a voluntary basis. The milking station 3 comprises an enclosure having an inlet gate 4 and an outlet gate 5, which are both capable of being opened automatically. The front end of the milking station 3 is denoted by 3a, the back end is denoted by

3b, the left side is denoted by 3c and the right side is denoted by 3d.

The milking station 3 comprises further an automatic milking machine (not explicitly illustrated) including teat cups 11 connected to an end unit by means of milk lines (only the portions attached to the teat cups 11 are shown in Fig. 1). The milking station further includes a milking robot or automatic handling device 14 having a robot arm 15 provided with a gripper. The milking robot 14 is arranged to automatically apply the teat cups 11 of the milking machine to the teats of a cow 8 present in the milking station 3 prior to milking. In Fig. 1 three of the teat cups 11 are arranged in a teat cup rack or magazine 16, whereas the fourth one is held by the gripper of the robot arm 15.

Typically, a teat cleaning device including e.g. a teat cleaning cup 21 or brushes 22 may be provided for cleaning the teats of the cow 8 prior to milking.

Further, the milking station 3 comprises an identification member provided to identify a cow approaching the milking station 3, and a central processing and control device 19, which is responsible for central processing and controlling of the animal arrangement, which inter alia includes the initiation of various activities in connection with the milking such as e.g. opening and closing of the gates 4 and 5, and control of the milking machine and its handling device 14. The central processing and control device 19 comprises typically a microcomputer, suitable software, and a database including information of each of the cows milked by the milking machine, such as e.g. when the respective cow was milked last time, when she was fed last time, her milk production, her health, etc.

A cow approaching the milking station is thus identified by the identification member, and the central processing and control device 19 may then, depending on the identification, give the cow access to the milking station 3 by means of opening the inlet gate 4. The teats may be cleaned by the teat cleaning device, after which the teat cups 19 are applied to the teats of the cow 8 in the milking station 9 under control of the central processing and control device 19.

The attachment is enabled by means of locating the teats of the cow 8 by a camera pair 23 directed towards the teats of the cow 8 when being located in the position to be milked. The camera pair 23 is provided to repeatedly record pairs of images. The central processing and control device 19 or other control or image processing device detects repeatedly the teats of the cow 8 and determines their positions by a stereoscopic calculation method based on the repeatedly recorded pairs of images. This position information is used by the central processing and control device 19 to send signals to the milking robot to move the robot arm to each teat after having gripped a respective teat cup.

Note that the robot arm 5 also has to move the teat cleaning cup 21 or the brushes 22 to the teats of the cow 8. This may be performed in the very same manner as the teat cups are moved.

During milking, milk is drawn from the teats of the cow 8 by means of vacuum being applied to the teat cups 11 via the milk lines, and the milk drawn is collected in the end unit. After the milking has been completed the teats of the cow may be subjected to after treatment, e.g. a spray of disinfectant, and then the outlet gate 5 is opened and the cow 8 may leave the milking station 3.

2. The camera arrangement

Where to place, i.e. where to arrange, and how to orientate, i.e. how to direct, the camera pair is an important feature. The task is to design an arrangement that should be capable of
5 detecting and locating every teat for each cow. Firstly, all teats must be seen by the cameras used. Secondly, the teats must be visualized in a manner that makes the image analysis easier. The implementation should also strive to maintain a low-cost profile in order to enable potential industrial production at
10 the same cost as the prior art systems used today.

In the preferred embodiment of Fig. 1, the camera pair 23 is mounted below the teats of the cow 8 and behind the cow when being located in the position to be milked so that the camera pair 23 is directed diagonally upwards towards the teats of the
15 cow 8 when being located in the position to be milked. The immediate position behind the cow 8 is in terms of image analysis a good location. The teats would be close to the cameras and by placing the cameras below the udders, directed diagonally upwards, both high and low udders can be detected in
20 most cases. A typical distance between the camera pair 23 and the teats of the cow 8 is 30 cm. A further advantage of this camera position is that the cameras will not be in the way of the robot arm 15 or the inlet gate 4.

The camera pair 23 may be movable up and down, depending on the
25 height of the udder, to attain a good contrast between the teats and background simply by moving the camera pair up or down.

In the preferred embodiment of Fig. 2, a first 23 and a second 24 camera pairs are directed towards the teats of the cow when being located in the position to be milked, wherein the each
30 camera pair 23, 24 is provided to repeatedly record pairs of

images, and an image processing device is provided for repeatedly detecting the teats of the cow and determining their positions by a stereoscopic calculation method based on the pairs of images repeatedly recorded by each of the camera pairs 23, 24. The first camera pair 23 is arranged behind the cow, whereas the second camera pair is arranged at the side of the cow. A typical distance between the second camera pair 24 and the teats of the cow is 80 cm.

At least two camera pairs may be necessary to obtain a high number of successful teat cup attachments. This is because the teats can occlude each other. More camera pairs may also relax the demands on the image processing.

The preferred embodiment of Fig. 3 differs from the embodiment of Fig. 2 in that the second camera pair 24 is mounted at a height so that the teats of the cow will belong to the outer contour of the cow in the repeatedly recorded pairs of images by that camera pair 24. Such measures would facilitate image processing considerably. The second camera pair 24 is preferably movable vertically in order to be positioned at a height so that the teats of the cow will belong to the outer contour of the cow in the repeatedly recorded pairs of images. Different cows have their udders at different heights and in order to secure that a teat contour is recorded the camera pair 24 should be movable in the vertical plane.

The preferred embodiments of Figs. 2-3 would not only improve the detection rate of the system but would also give more reliable stereo calculations since distance calculations from two independent systems would be used.

The movable cameras could be trained for each cow or be trained the first time a cow enters the milking station 3. The result

could be stored in a database and be updated after each milking. Thus, by having the possibility to move the cameras of the camera pairs relative the milking station 3 and the cow when being located in a position to be milked, and optionally relative each other, the number of failed teat position determinations due to teats being obscured can be reduced.

The preferred embodiment of Fig. 4 differs from the embodiment of Fig. 3 in that each of the two camera pairs, here denoted by 41 and 42, comprises three cameras 41a-c and 42a-c. By such a stereo vision system three stereoscopic position calculations can be made for each camera pair 41, 42, and thus in total six calculations can be made provided that each camera 41a-c and 42a-c can see the teat in question. It would of course be even better to use even more camera pairs to achieve even further stereoscopic position calculations, or at least render it most likely that two cameras in one camera pair can always detect a teat of a cow.

According to yet a preferred embodiment of the invention, the camera pair is located straight below the cow's udder, and the cameras are directed upwards. For instance the camera pair may be located in the floor of the milking station behind a translucent protecting window provided with cleaning means such as wipers. The camera pair together with the protecting window could be raised by a telescopic arrangement to be positioned closer to the teats of the cow. Such camera pair would always be capable of detecting all four teats of the cow. However, the position is extremely vulnerable for dirt and damages.

It shall be noted that the cameras of each camera pair in the preferred embodiments described above are advantageously arranged vertically one above the other. The image processing device may then, for each teat and for each pair of images,

define the position of the lower tip of the teat contour in the pair of images as conjugate points for the stereoscopic calculation, and find the conjugate points along a substantially vertical epipolar line. This increases the accuracy and precision of the teat position determination considerably. This is discussed in detail in section 5 with reference to Figs. 6 and 7.

Further, the cameras of each camera pair are arranged so that the image planes of the cameras of each camera pair are coplanar. This is not a prerequisite for the stereoscopic calculations, but it facilitates the same.

The milking station 3 is typically arranged in a barn. Neither the choice of location nor the indoor/outdoor lighting should affect the performance of the stereo vision system. Lightning is an issue of both suppressing natural light differences and enhancing the image quality to facilitate image analysis.

Therefore, an artificial light source particularly provided for illuminating the udder of the cow to thereby increase the contrast in the repeatedly recorded pairs of images is preferably used.

In the preferred embodiment of Fig. 2, a light source 25 is provided, which may emit white light, colored light or UV light. In order to obtain a good contrast between the teats and the udder or other kind of background, the light source 25 ought to create a back lighting or sidelight for each camera pair. In Fig. 2 the light source 25 is arranged at a low position at the left side 3c of the milking station, but further ahead of the second camera pair 24. The light is directed backwards and diagonally upwards toward the teats of the cow. Thus, a back lighting is obtained for the first camera

pair 23 and a sidelight is obtained for the second camera pair 24.

The light source 25 could be movable and capable of being directed toward different directions in order to find an appropriate illumination of the teats of the cow.

An external light source for the particular purpose of illuminating the teats of the cow is preferred but not compulsory. However, if such light source is not used image analysis has to be used to counteract various lightning effects. Methods that can be used include e.g. histogram equalization, a method used to divide the grey levels equally within the entire range. However, since both color and texture of the teats are similar to the surface of the udder the contrast between them may in some instance be too low.

15 3. Image processing: teat detection

Fundamental steps in image processing include the following different processing steps.

- 20 > Pre-processing
- > Segmentation
- > Representation and description
- > Classification

25 The first step is pre-processing. This stage aims to enhance the image by removing noise or distortions. The main idea is to increase the chances to have success in the later processing steps by apply methods to suppress some data and to enhance other. Common techniques are transformations for enhanced

contrast and noise removal. The transformations can be combinations of smoothing, different edge detectors or logical operations.

5 Next, there is a group of processing techniques commonly referred to as segmentation. These techniques intend to divide a digital image into objects, i.e. "interesting areas", and background.

10 Once the image is divided into objects and background the objects need to be described using descriptors. These descriptors could be size, grey levels, roundness etc. Simultaneously as the descriptors are calculated each object is given a unique identification number called a label. This step is referred to as representation and description.

15 Finally, a classification algorithm is applied to the object list. This is a way to use object information to associate the objects with objects in the milking station which have been imaged by the camera pair. Typically, some pre-knowledge of the images recorded is needed to create a good classification algorithm. The pre-knowledge is important in image analysis, not
20 just for the classification step. The more information available regarding a specific type of images, the better the image analysis can be made.

25 In order to accurately detect teats, an automatic method has to be created. The accuracy of the detection system should preferably be no less than prior art systems of today. To solve such a complex problem it is necessary to divide it into smaller processing steps described above. Fig. 5 shows various methods that can be used for teat detection and to which of the processing steps they belong. The most preferred methods are
30 indicated in Fig. 5 and outlined below.

Mean filtering is a blurring pre-processing method, which softens sharp areas in an image. If used properly it could remove noise from the image and potentially improve the image before it is processed by other methods. The method uses the spatial relationship of pixels to smooth an image. A filter mask of size $M \times N$ pixels is applied. A pixel, usually the center pixel, in that mask achieves a new pixel value, which represents the mean value of the pixel values within the mask. The bigger the filter mask the more the image is blurred.

Motion detection is a pre-processing method for removal of superfluous information in the images by utilizing the cow's movement. Two images are recorded after one another, with a small time difference, the fact that the grey level values will have changed were movement has occurred is utilized in the method. Since a cow never is stationary, because of breathing or balance correction the time between the two consecutive images can be small.

A black and white image is created, where white pixels represent important areas (i.e. areas where the cow's udder likely is) and black pixels represent uninteresting background areas. This is performed by changing pixels to white in areas where the absolute value of $\text{pixel}_{\text{image1}} - \text{pixel}_{\text{image2}}$ exceeds a given constant and every other pixel is changed to black. Such a template image will be used to extract the interesting areas in the original image. Since the binary template image is scattered, it needs further processing to be useful. A good manner to remove scatter is to relax the image. A relaxation method gives the center pixel, in a neighborhood, the value of the most occurring value in that neighborhood. The larger the neighborhood is, the coarser the image becomes.

Since it is critical to include the udder area in the final image it is a good idea to increase the white area in the template by a pixel expanding algorithm. This could be done by dividing the image into $K \times L$ squares, with an actual size of $M \times N$ and changing the pixels of the entire area to white if there is any occurrence of white pixels inside the area.

If the teats are located in the outer part of the cow's contour, all other remaining information about the cow is here superfluous. This information is removed by applying a method, which is called *contour creation*. Contour creation firstly copies the black and white template image and then performs a number of binary erosions on the copy. The result is a smaller copy of the original, where every white area has shrunk. By subtracting the shrunk image from the original template image, the result is an image with white bands around the contour of the cow.

By combining this *band template image* with the originally recorded image and exchanging every white pixel with its corresponding grey pixel, an image with a band of grey levels is created. This image is further processed.

This last step can typically not be applied fully to images taken from the rear position since the teats in those images are not part of the outer contour of the cow. By only using the first three steps of the motion detection method background areas are removed from the image, which still reduces the amount of data considerably.

Thus, according to the present invention, an image processing device is provided to perform any of the above mentioned steps to reduce the area in which the teats likely are.

Experiments have shown that the phase plays an important role in the perception of visual features. The most common features in images consist of combinations of steps, roofs and ramp profiles as well as mach band effects. There is no single linear feature detector that can detect those combinations. But by using the phase component of the Fourier Transform these features can be extracted. The phase plays a decisive role in the perception of visual features. If a human was asked to draw a sketch of the image, localizing precisely the edges or markings of interest as seen in the scene, then chosen points would be those where there is maximal order in the phase components of a frequency-based representation of the signal. Unfortunately, the phase congruency model is hard to implement but could be approximated with local energy in a signal. So instead of searching for local maxima in the phase congruency function, local maxima in the local energy function should be found.

The outputs of this method are two images, one containing magnitude and one orientation of the features. So the phase congruency feature detector is therefore suitable to be used in combination with a canny edge detection algorithm, see below. One issue that needs to be handled is that in combination with the motion detection method erroneous edges will be created where data has been removed. Since removed data is represented as black sections within the image, edges will occur in the contact area between the black edge and the important data area.

Thus, according to the present invention, an image processing device is provided, for each time the teats are to be detected, to apply an edge detection algorithm based on the phase congruency model of feature detection to thereby find edges and corners in a recorded pair of images that most likely include those of the teats of the cow.

A preferred edge-detection method is a Canny detection algorithm.

Next, a representation and description processing method is applied for calculating features of the edges and corners found in the recorded pair of images. Preferably, a labeling algorithm, such as a connected-components labeling method, and a feature vector method are used.

Finally, a classification method is applied for identifying edges and corners of the found edges and corners in the recorded pair of images which belong to the teats of the cow based on the calculated features. The classification method includes typically a low-level classification algorithm and a high-level classification algorithm such as a hierarchical chamfer matching algorithm.

15 4. Image processing: calibration

If stereo vision is to be useful, calibration is a decisive component. Stereo vision or stereoscopic calculation strongly relies on accuracy in calibration and measurements.

The basic idea is built on the ideal case: a pinhole camera that is perfectly aligned. This is not the case in reality. To calculate scene point coordinates, additional information is needed. There are the camera constant and the baseline distance. In reality the relative position of the cameras differs more than just the baseline displacement. The camera also has other important parameters other than the camera constant. The image center can be displaced and there are probably many types of distortions in the image. There are for example radial distortion and de-centering distortion. Radial distortion is seen as a scene point being imaged closer or further out from

the image centre than it should have been. This distortion is dependent on the distance from the image center.

5 The camera intrinsic parameters are collected with the interior calibration. This gives the actual image center and other distortions parameters. These are calculated to be able to reconstruct the actual image.

10 Before the resulting images can be used and distance calculations can be done, the relation between the camera positions has to be determined. This is achieved by relative calibration. Here both parameters that describe the rotation and displacement are estimated.

15 If the calculated distance to a scene point should be useful, its relation to another coordinate system is significant. The absolute calibration is used to determine this relation, both rotation and transformation, between two coordinate systems. In this case the relation between a stereo vision system and the robot coordinate system is of interest.

20 It has been found experimentally that absolute calibration is much more difficult than relative calibration, and that much more reference points in the milking station are needed. Further, the stereo vision system must be continuously calibrated due to movements and other disturbances, which do occur in the milking station, and which affects calibration points placed therein. The accuracy of the calibration points must be 10 times higher than the accuracy demand on the teat positioning. The parameters needed in the stereo calculation, is the position and rotation of both cameras in the milking station coordinate system, the image center and the radial and de-centering distortion factors.

Due to the difficulties obtained when performing absolute calibration to obtain highly accurate absolute teat positions, the present invention proposes a manner of operation, which remedies this problem.

5 An image processing device is provided in an initial stage of a teat cup attachment to roughly determine the absolute position of a teat of the cow in a coordinate system of the milking station, by aid of which the robot may fetch a teat cup and move
10 it to a position close to a teat of the cow present in the milking station. The image processing device is then provided, i.e. in a later stage of the teat cup attachment when the robot arm is close to attach the teat cup to the teat of the cow, to repeatedly and accurately determine the position of the teat of the cow relative the robot arm or the teat cup. Thus, the
15 position of the teat of the cow relative the robot arm or the teat cup is more exact than what the absolute position can be measured.

The image processing device is provided, in the later stage when the robot arm is close to attach the teat cup to the teat of the
20 cow, to repeatedly determine the relative position of the teat of the cow in a coordinate system of the camera pair used, and to repeatedly detect the robot arm or the teat cup and determine its relative position in the coordinate system of the camera pair by the stereoscopic calculation method of the present
25 invention.

During each absolute calibration, the camera pair is provided to record a pair of images, wherein several well defined points are located in the common field of vision of the camera pair, the positions of the well defined points being known in the
30 coordinate system of the milking station, and the image processing device is provided to perform an absolute

calibration process, in which the positions of the image planes of the cameras of the camera pair are determined in the coordinate system of the milking system to thereby be capable of determining the absolute position of the teat of the cow.

- 5 Similarly, during each relative calibration, the camera pair is provided to record a pair of images, wherein several well defined points are located in the common field of vision of the camera pair, and the image processing device is provided to perform a relative calibration process, in which the positions
10 of the image planes of the cameras of the camera pair are determined relative each other to thereby be capable of determining the position of the teat of the cow relative another measured position.

- Advantageous by this process is that a very fast and accurate
15 teat cup attachment is enabled, while no extremely high demands are put on the calibration process.

5. Image processing: stereo calculation

- Once the teats of the cow have been detected their positions have to be determined by a stereo vision or stereoscopic calculation method. Stereo vision is the recovery of three-
20 dimensional information of a scene from multiple images of the same scene. Stereo vision in computer systems tries to copy the human way when calculating distance. The principle is based on triangulation. Points on the surface of objects are imaged in
25 different relative positions in the images recorded depending on their distance from the viewing system. The basic model is two ideal pinhole cameras viewing the same scene, only with a baseline distance separating them. The image plane should preferably be coplanar, which means that there is no relative
30 rotation between the cameras.

In order to compute the best baseline distance for a system a number of parameters are to be considered. Camera positions, type of cameras, image area, relative positions, milking station constraints and distance-accuracy demands. The baseline is therefore a trade-off of the mentioned parameters. Measurements have shown that a baseline distance of about 5-20 cm would be a suitable choice. Further investigations have shown that the baseline distance should be about $1/5$ of the distance from the cameras to the object of interest. The camera pair placed on the side of the milking station would then have a baseline distance of approximately $80/5 \text{ cm} = 16 \text{ cm}$. The rear camera pair behind the cow is closer to the object and should have a baseline distance of about $30/5 \text{ cm} = 6 \text{ cm}$. If the disparity increases the probability of finding and detecting the conjugate pair decreases.

One scene point is imaged in different locations in a first image and a second image recorded by differently placed cameras. The scene point in the first image plane and in the second image plane are both placed on the epipolar line and is called a conjugate pair or conjugate point. The epipolar line is the line that connects the scene point in the first and second images. If the scene point is found in the first image, the scene point in the second image is on the epipolar line. In the ideal model this means that if the cameras are arranged horizontally side by side the first and second image points are on the same image row. Conjugate points are a pair of image coordinates, one from each image, which are each a projection from the same point in space onto the respective image plane. The displacement between the image points is called the disparity and will finally give the depth information.

For the stereo vision calculations the choice of the points that constitutes the conjugate pair is important. They have to represent interesting measuring points and should not result bring any additional error to the distance errors. The
5 conjugate points should also be easy to find.

According to the invention the position of the lower tip of the teat contour in the pair of images is defined as conjugate points.

The usual manner to model a stereo vision system is to have a
10 horizontal displacement of the cameras. Fig. 6 shows a pair of images 91a-b as taken by a camera pair mounted horizontally side by side. The contour images of the teats are denoted 92a and 93a in the left hand image 91a, and 92b and 93b in the right hand image 91b. The epipolar lines are denoted by 94 for
15 the teat images 92a-b and by 95 for the teat images 93a-b. This would result in difficulties to find the conjugate points because the lines representing the lower tips of the teat contours are essentially horizontal and the epipolar lines are horizontal. Since these lines are almost parallel it is
20 difficult to select a correct conjugate point.

According to a preferred embodiment of the present invention, the cameras of each camera pair are arranged vertically one above the other as has been disclosed in section 2 above. Fig.
25 7 shows a pair of images 101a-b as taken by a camera pair mounted vertically one above the other. The contour images of the teats are denoted 102a and 103a in the upper image 101a, and 102b and 103b in the lower image 91b. The epipolar lines are denoted by 104 for the teat images 102a-b and by 105 for
30 the teat images 103a-b. This results in difficulties to find the conjugate points because the lines representing the lower tips of the teat contours are essentially horizontal and the

epipolar lines are horizontal. Since these lines are almost parallel it is difficult to select a correct conjugate point.

Through this arrangement of the cameras, the epipolar lines become orthogonal to the lines representing the lower tips of the teat contours. This will increase the accuracy and simplifies the conjugate point detection.

Additional information or conjugate points may be needed. This can be achieved by selecting further points on the lines representing the lower tips of the teat contours. The teats internal orientation will indicate on which teat line the conjugate point is to be found. The thickness and teat angle are two interesting parameters. These parameters should not be difficult to extract from the teat tip lines.

In stereovision systems different kinds of occlusion occur. Depending on the cameras relative positions the two images from a stereo-camera pair differ. One teat can be visible in one image but be partly or fully occluded in the other stereo-image pair. Another issue is the fact that edges from a teat visualized in two images, are in fact not the same physical edge.

Fig. 8 is a side view of the lower part of a teat of a milking animal illustrating self-occlusion. The teat tip 111 is modeled as a half sphere. Two projection lines 112a and 112b from cameras at different heights illustrate the fact that the lower tip of the teat contour in the images recorded by the cameras at different height does not correspond to the same physical scene point on the teat of the cow. In the image by the upper camera point 113a will be recorded as the lower tip of the teat contour, whereas in the image by the lower camera point 113b will be recorded as the lower tip of the teat contour. This

leads inevitably to errors in the calculation of the positions of the teats.

5 The phenomenon which may be compensated for by still a preferred embodiment of the invention. An image processing device can be provided, for each teat and for each pair of images, to compensate for any deviations caused by the fact that the lower tip of the teat contour in the pair of images corresponds to different objection or scene points on the teat due to different perspective views, in which the pair of images are recorded, by means of creating a mathematical model of the
10 characteristic form of the teat, and to calculate the compensation based on the mathematical model, the different perspectives, and the distance to the teat.

15 The exact distance to the teat is typically unknown, whereupon the image processing device can be provided to determine the positions of the teats of the milking animal, and to calculate the compensation for any deviations caused by the fact that the lower tip of the teat contour in the pair of images corresponds to different objection points on the teat, in an iterative
20 process, wherein better teat positions and better compensation can be obtained for each iteration.

6. Image processing: further functions

When using a stereo vision camera system it is possible to expand the functionality to more than locating teats and robot
25 arms or teat cups. For instance, an image processing device can be provided to automatically detect injuries and/or dirt on the teats of the cow by an image processing method based on one or several of the repeatedly recorded pairs of images. A single image could be used to discover visual defects, such as redness, cuts, sores or wounds, on the teats and the udders of the cows.
30

The camera system may also surveillance the milking process or inspect almost any part of the milking station. This would be easier if the cameras could be rotated independently of each other.

5 Furthermore, the teat cleaning device of the milking station may be designed to be capable of cleaning teats of cows according to anyone of a plurality of different teat cleaning schemes, and then one of the plurality of different teat
10 cleaning schemes can be selected for the cleaning of the teats of the cow based on the automatic detection of injuries and/or dirt on the teats of the cow. Teat cleaning can in this manner not only be performed on a cow individual basis but also on a teat individual basis.

15 The stereo vision system can also be used to create 3D images of objects. In those images the teat thickness and deformations can be analyzed.

Further, the stereo vision system can be used when a technician monitors the milking station, or inspects the milking station for errors. Thus, the stereo vision system can operate as a
20 simple surveillance system. It can easily detect any movement around the robot and trespassers could be detected.

The stereo vision system is vulnerable to dirt or damages on the lenses of the cameras. By using image information dirt could be detected and allow for the system to attend to the problem.

25 By recording an image of the milking station e.g. when it is empty, a good reference image is created. The image can then be compared with later images. If the images do not differ more than to a given extent, no dirt has settled on the lens. This method is suitable for detecting lumps of dirt but it will

probably not work well when it comes to detect slowly smeared lenses. Since this kind of dirt seldom appears suddenly, but is rather a slow deterioration in image quality. However, a smeared image could be compared to a strong blurring filter: the number of edges found in an image would be smaller the blurrier the lens is. Thus, by using images of the empty milking station it is probably possible to discover lumps of dirt and smeared areas by applying two different techniques.

If the cameras can be rotated around their own axes it is also possible to rotate them into such positions where they could detect dirt on each other. A lens is an easy object to automatically detect and it is possible to create a method for detecting dirt on an otherwise very clean surface.

11
12
13
14
15
16
17
18
19
20

CLAIMS

1. An arrangement for determining positions of the teats of a milking animal in a milking system comprising a robot arm for automatically attaching teat cups to the teats of a milking animal when being located in a position to be milked, and a control device for controlling the movement of said robot arm based on determined positions of the teats of the milking animal, said arrangement comprising:
- a first camera pair directed towards the teats of the milking animal when being located in the position to be milked, wherein the first camera pair is provided to repeatedly record pairs of images;
 - an image processing device provided for repeatedly detecting the teats of the milking animal and determining their positions by a stereoscopic calculation method based on said repeatedly recorded pairs of images, characterized in that
 - said image processing device is provided to automatically detect a condition of the teats of the milking animal by an image processing method based on said repeatedly recorded pairs of images.
2. The arrangement of claim 1 wherein said image processing device is provided to automatically detect injuries and/or dirt on the teats of the milking animal by an image processing method based on said repeatedly recorded pairs of images.
3. The arrangement of claim 2 wherein said injuries are any of wounds, cuts, sores, or red spots.
4. The arrangement of claim 1 wherein said image processing device is provided to automatically detect if a teat of the

milking animal is swell up by an image processing method based on said repeatedly recorded pairs of images.

5 5. The arrangement of any of claims 1-4 wherein said milking system comprises at least one teat cleaning device capable of cleaning teats of milking animals according to anyone of a plurality of different teat cleaning schemes, said arrangement comprising means provided to select one of said plurality of different teat cleaning schemes for the cleaning of the teats of said milking animal based on said automatic detection of the
10 condition of the teats of said milking animal.

6. The arrangement of any of claims 1-5 wherein said image processing device is provided to automatically detect dirt or damages on the lenses of said first camera pair by an image processing method based on repeatedly recorded pairs of images
15 by said first camera pair.

7. The arrangement of claim 6 wherein said image processing device is provided, for each one of the cameras of said first camera pair, to automatically detect dirt or damages on the lens of that camera by an image processing method based on
20 comparisons of repeatedly recorded images by that camera.

8. The arrangement of claim 6 or 7 wherein

- the cameras of said first camera pair are rotatable to be capable of being directed towards each other; and

25 - said image processing device is provided, for each one of the cameras of said first camera pair, to automatically detect dirt or damages on the lens of that camera by an image processing method based on repeatedly recorded images by the other camera.

9. A method for determining positions of the teats of a milking animal in a milking system comprising a robot arm for automatically attaching teat cups to the teats of a milking animal when being located in a position to be milked, and a control device for controlling the movement of said robot arm based on determined positions of the teats of the milking animal, said method being characterized by the steps of:

- directing a first camera pair towards the teats of the milking animal when being located in the position to be milked;
- repeatedly recording pairs of images by the first camera pair;
- repeatedly detecting the teats of the milking animal and determining their positions by a stereoscopic calculation method based on said repeatedly recorded pairs of images, wherein
- injuries and/or dirt on the teats of the milking animal are automatically detected by an image processing method based on said repeatedly recorded pairs of images.

10. An arrangement for determining positions of the teats of a milking animal in a milking system comprising a robot arm for automatically attaching teat cups to the teats of a milking animal when being located in a position to be milked, and a control device for controlling the movement of said robot arm based on determined positions of the teats of the milking animal, said arrangement comprising:

- a first camera pair directed towards the teats of the milking animal when being located in the position to be milked, wherein

the first camera pair is provided to repeatedly record pairs of images;

- an image processing device provided for repeatedly detecting the teats of the milking animal and determining their positions by a stereoscopic calculation method based on said repeatedly recorded pairs of images, characterized in that

- said image processing device is provided to automatically detect dirt or damages on the lenses of said first camera pair by an image processing method based on repeatedly recorded pairs of images by said first camera pair.

11. The arrangement of claim 10 wherein said image processing device is provided, for each one of the cameras of said first camera pair, to automatically detect dirt or damages on the lens of that camera by an image processing method based on comparisons of repeatedly recorded images by that camera.

12. The arrangement of claim 10 or 11 wherein

- the cameras of said first camera pair are rotatable to be capable of being directed towards each other; and

- said image processing device is provided, for each one of the cameras of said first camera pair, to automatically detect dirt or damages on the lens of that camera by an image processing method based on repeatedly recorded images by the other camera.

13. A method for determining positions of the teats of a milking animal in a milking system comprising a robot arm for automatically attaching teat cups to the teats of a milking animal when being located in a position to be milked, and a control device for controlling the movement of said robot arm based on determined positions of the teats of the milking

animal, said method being characterized by the steps of:

- directing a first camera pair towards the teats of the milking animal when being located in the position to be milked;

5 - repeatedly recording pairs of images by the first camera pair;

- repeatedly detecting the teats of the milking animal and determining their positions by a stereoscopic calculation method based on said repeatedly recorded pairs of images,

10 wherein

- dirt or damages on the lenses of said first camera pair is/are automatically detected by an image processing method based on repeatedly recorded pairs of images by said first camera pair.

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918

1/4

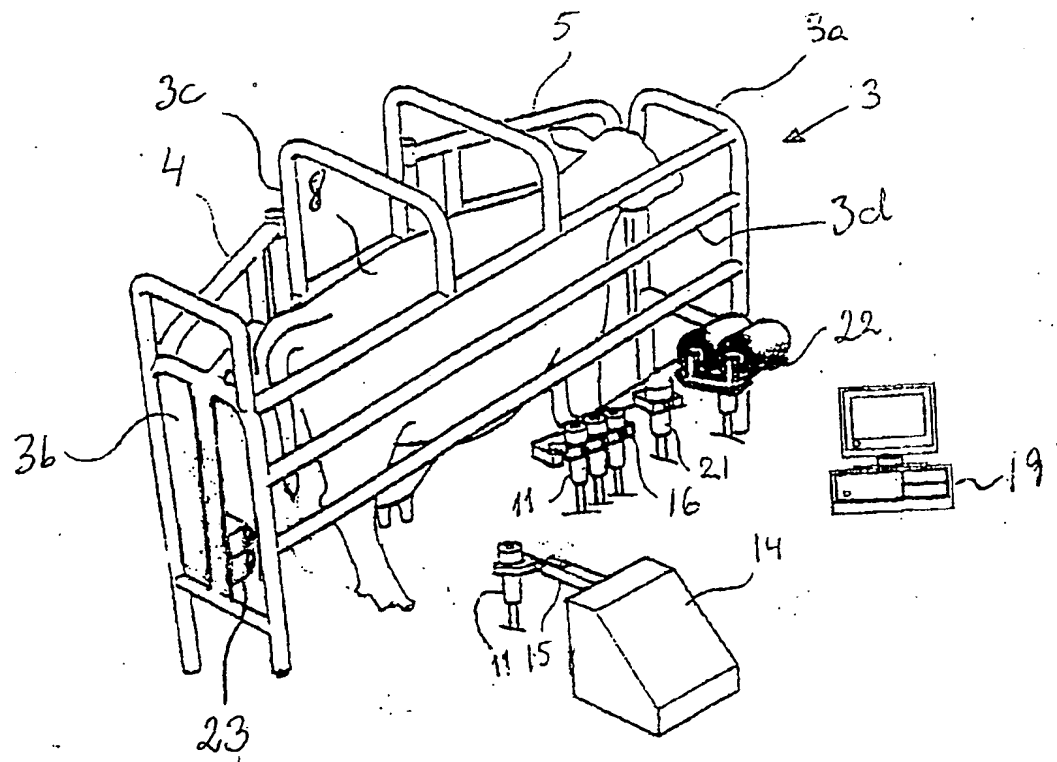


Fig. 1

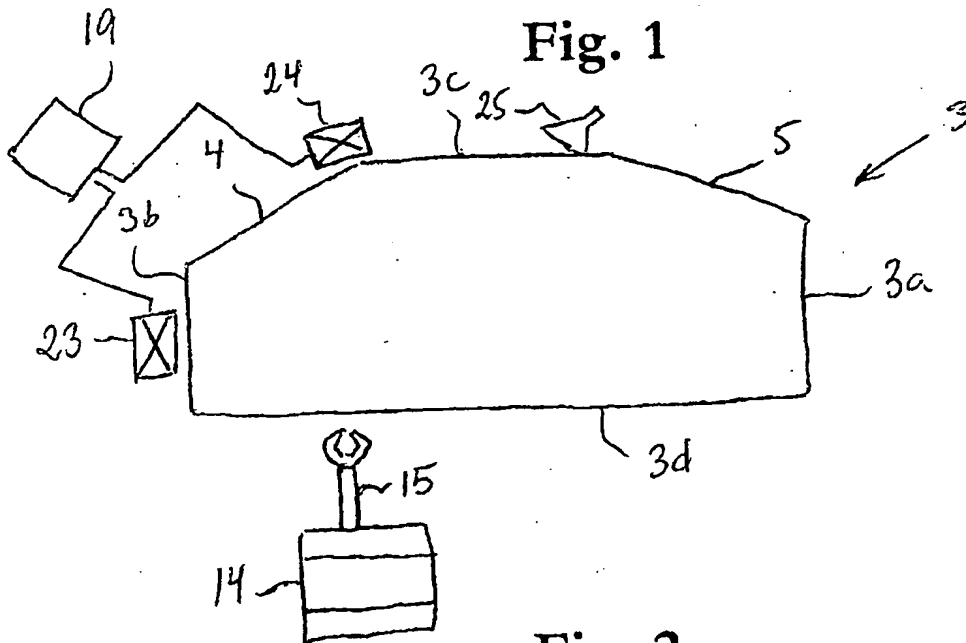


Fig. 2

2/4

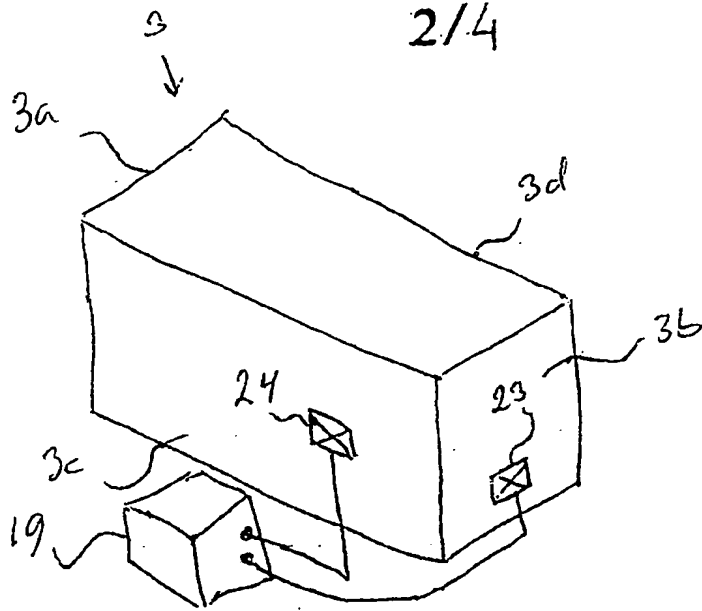


Fig. 3

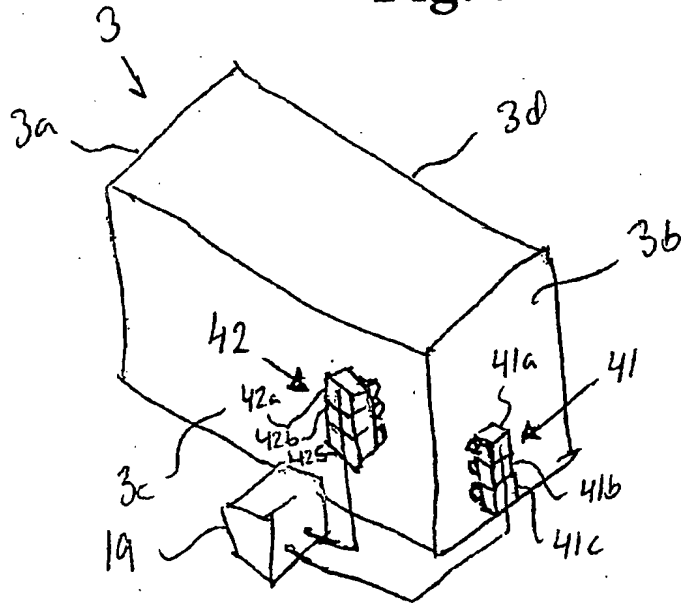


Fig. 4

3/4

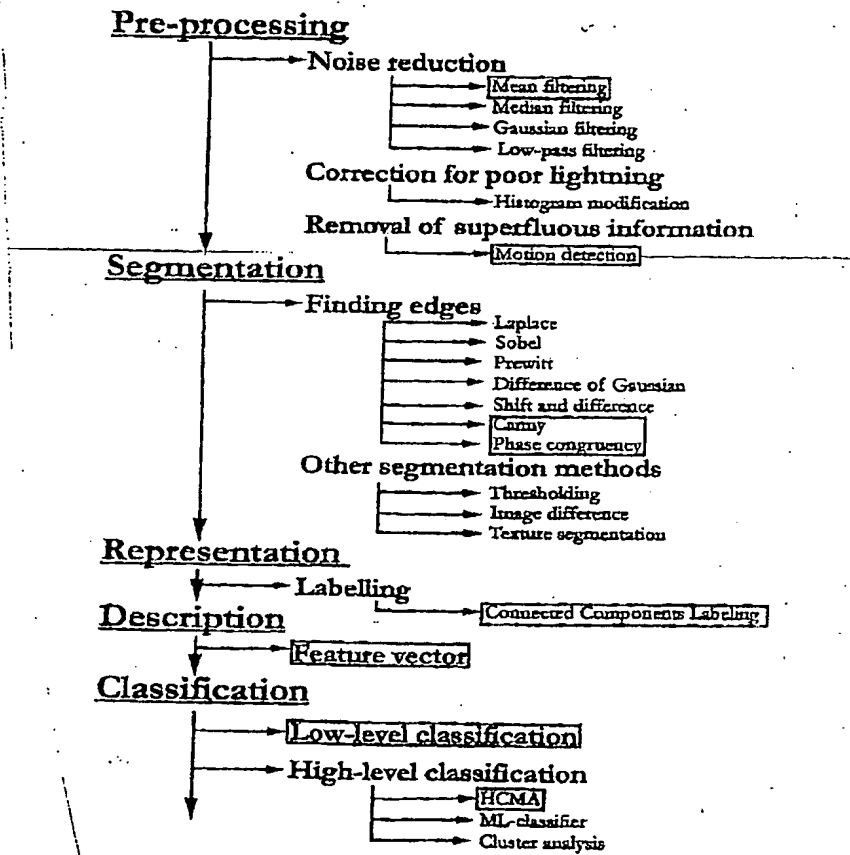


Fig. 5

4/14

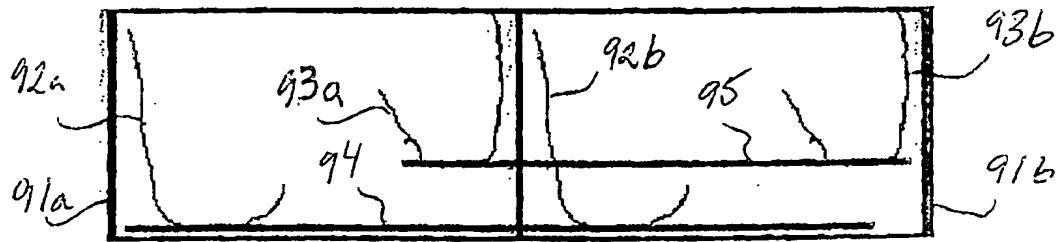


Fig. 6

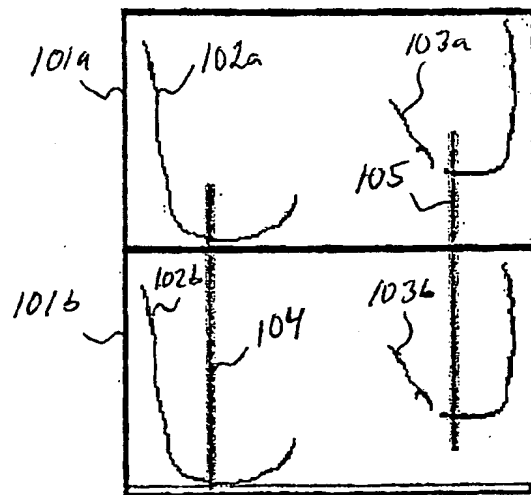


Fig. 7

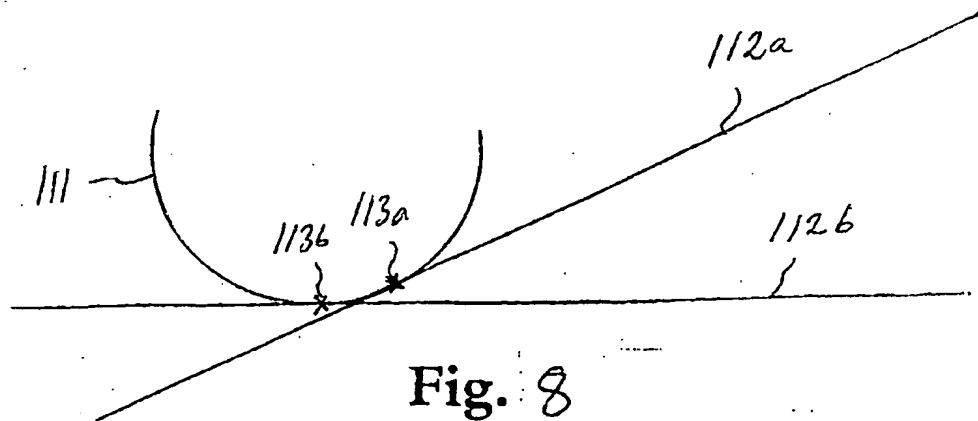


Fig. 8